README File for 2020 Census Demographic and Housing Characteristics Summary File (v. 2023-05-25) Delivered via File Transfer Protocol (FTP)

Note: *Users processing these FTP files in a Windows environment should carefully read the File Information section of this document.*

Contents

About the FTP Application
The FTP Directory
File Information
Segmented Data
File Record Layout
Other Sources of the Data
Clearance for Public Release

About the File Transfer Protocol (FTP) Application

This FTP application is intended for experienced users of census data, zip files, and analytical/database software. Due to the size of the files, the FTP user should have a fast file transfer capability.

The FTP Directory

The FTP directory is located at <https://www2.census.gov/programs-surveys/decennial/2020/data/demographic-and-housing-characteristics-file/>.

The data are organized in zip files, for the National Summary File, and by state, District of Columbia, and Puerto Rico for the State Summary Files. The National and State Summary File data are included in the same directory. When the data are added to each respective subdirectory, the subdirectory will contain a single zip package that includes a geographic header file and a number of data segment files. Refer to the Segmented Data section below for more information.

Users of the FTP site need to unzip the package after downloading, then import the data into spreadsheets or analytical/database software of their choice for data analysis and table presentation. We are unable to provide one-on-one support for applications of the data to specific

spreadsheets or analytical/database software requirements.

File Information

Once the package is unzipped, the files are in the Unicode Transformation Format (UTF-8). The geographic header file and data files are pipe delimited. Both the geographic header and the data files contain geographic linking fields. These files have been constructed in a LINUX environment.

For successful use with many programs running in a Windows environment, these files need to be modified to use the UTF-8 carriage return/linefeed sequence, chr(13) + chr(10) as a record terminator. This is an easy step in the UnZIP process using any UnZIP software which offers the conversion option.

The resulting file will meet the ANSI MS-DOS/Windows standard used by Access and other MS Windows-based programs. If the data are being processed in a LINUX environment, they can be unzipped using any standard Linux ZIP/UnZIP package.

Segmented Data

The data in the 2020 Census Demographic and Housing Characteristics Summary File are segmented. This is done so that individual files will not have more than 255 fields in most cases, facilitating exporting into spreadsheets or analytical/database software. In short, **to get a complete set of a state's 2020 Census Demographic and Housing Characteristics Summary File, users must use all files in the package.**

Each zip Summary File contains a geographic header file and a number of segmented files described in the Table Matrix section in the technical document of the data product.

A unique logical record number (LOGRECNO in the geographic header) is assigned to all files for a specific geographic entity; all records for that entity can be linked together across files. Additional identifying fields are also carried over from the geographic header file to the table files. These are file identification (FILEID), state/U.S. abbreviation (STUSAB), characteristic iteration (CHARITER), and characteristic iteration file sequence number (CIFSN), which is the segment number.

The geographic header file contains the geographic codes and other fields that identify the specific geographic entities that are linked to the table files. The geographic header record layout is identical across all electronic data products from the 2020 Census. Note that some fields are not filled. For example, the characteristic iteration (CHARITER) field will be used in later 2020 Census data products but in the 2020 Census Demographic and Housing Unit Characteristics Summary File, it is always coded as 000.

File Record Layout

For a layout of the individual tables for each file, refer to Chapter 3, Data Dictionary, in 2020census-demographic-and-housing-characteristics-file-and-demographic-profile-techdoc.pdf <https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/demographic-and-housing-characteristics-file-and-demographic-profile/2020census-demographic-and-housing-characteristics-file-and-demographic-profile-techdoc.pdf>.

Other Sources of the Data

In addition to the 2020 Census Demographic and Housing Characteristics Summary File being available through the Census Bureau FTP site, the 2020 Census Demographic and Housing Characteristics Summary File data will be available primarily through the Census data exploration platform, data.census.gov, by May 25, 2023.

To request assistance with these products, contact the Census Bureau: <mailto:2020DAS@census.gov>

Clearance for Public Release

Public release of the 2020 Census Demographic and Housing Characteristics File (DHC) (v. 2023-05-25) has been authorized by the Census Bureau's Disclosure Review Board (clearance number CBDRB-FY21-DSEP-005).